**Online Appendix:**

# Excess Deaths In the United States During the First Year of COVID-19: Supplemental Methods and Results

## Supplemental Methods and Results

### Original Analysis of Excess Deaths

*Data*

Information on monthly death counts from March 2009 through December 2020 were obtained from the Centers for Disease Control (CDC) and Prevention *Multiple Cause of Death (MCOD)* data, downloaded on May 14, 2022 from *CDC Wonder* (1), which provide information from death certificates on a single underlying cause of death (UCD), up to 20 additional causes, and demographic variables. Data were utilized on cause of death, using four-digit *International Classification of Diseases, Tenth Revision* (ICD-10) codes, age, race/ethnicity, gender, education, and year of death. These data cover the universe of fatalities occurring to residents of the United States from 1999-2019. In the public-use data, downloadable from *CDC Wonder*, subnational data with fewer than 10 persons are suppressed. However, data suppression is not an issue for this analysis because suppression does not occur at the national level. Since final death counts from January 2021 through February 2022 were not available at the time of this study, provisional data for these months were also obtained from *CDC Wonder* (2). These data were updated on May 1, 2022 and downloaded on May 14-15, 2022.

Using these data, total mortality counts were computed as were those for selected causes and by sex, race/ethnicity, and age. For brevity, non-Hispanic whites, blacks, and other races are referred to below as whites, blacks, and other races. The following causes of death were examined: COVID-19; diseases of the heart (heart); malignant neoplasms (cancer); cerebrovascular disease (stroke); Alzheimer's disease; diabetes mellitus (diabetes); chronic lower respiratory disease (respiratory); influenza or pneumonia (flu/pneumonia); nephritis, nephrotic syndrome or nephrosis (kidney); motor vehicle accidents (vehicle); drug overdoses (drug); intentional self-harm (suicide); and assault (homicide). ICD-10 codes for these causes of death are shown in Table A1.

The provisional death data may have been incomplete or reported with delays. These issues are expected to be minor since the data were updated on May 1 2022, a full 14 months after the last month of data used (February 2021). However, there may also be other issues. Coronavirus disease deaths were reported if COVID-19 contributed to the death, whether or not it was the underlying cause, but this classification might have been incomplete, particularly early in the pandemic when COVID-19 may have been inconsistently reported. There is also some inconsistency in the reporting of race/ethnicity between the final and

provisional data. Specifically, in the final data (covering the period through December 2020), there are race categories for: White, Black or African American, American Indian or Alaska Native, and Asian or Pacific Islander. The other race category includes non-Hispanics who are American Indian/Alaska Natives or Asian/Pacific Islanders. However, in the provisional data the categories are: Whites, Blacks or African American, Asian, Native Hawaiian or Other Pacific Islander, American Indian or Alaska Native, and More than One Race. For these data, the other race category includes non-Hispanics in all of these except for the white or Black/African American groups.

*Actual and Baseline Death Counts*

For simplicity, throughout the remainder of the discussion, each March through February period will often be labeled as a "year" and denoted by the year with the majority of these months. For example, 2012 refers to March 2012 through February 2013 and 2020 indicates March 2020 to February 2021.

Observed fatality counts for all deaths or for the specified group or cause were calculated for each month between March 2009 and February 2021. Adjustments were made for the extra day in February during the leap years of 2012, 2016, and 2020 by multiplying February deaths in these years by 28/29. The data were then aggregated over the 12-month March to February "years" to provide observed mortality counts, $D_{it}$, for all deaths or specified group or cause $i$ in year $t$.

Year-over-year percentage changes were computed as:

$$\Delta D_{it} = \frac{D_{it} - D_{it-1}}{D_{it-1}} \times 100\%, \tag{A.1}$$

for $t = 2010 - 2019$.

Changes in monthly total death counts, compared to the same month in the prior year, were also calculated and these revealed substantial excess idiosyncratic volatility in January, relative to other months. For instance, the standard deviation in the year-over-year January change was more than twice as large as the next highest month (December) and almost five times as big as for the average of the eleven non-January months (see Table A2). This was partly due to the higher average number of deaths in January than other months but mostly for other reasons. This can be seen by noting that the coefficient of variation of the year-over-year January change is 4.3 times the average for the other eleven months and 1.9 times that of December, the month with the next highest coefficient of variation. This excess volatility, which preliminary work indicated was due to a variety of causes including, importantly but not exclusively, the severity of seasonal influenza/pneumonia epidemics,

increased the difficulty in forecasting baseline deaths (see the lower panel of Table A2). Therefore, these were calculated based on an adjustment to mortality counts described next.

Define deaths in March through December plus February (i.e. annual counts without January deaths) in year $t$ by $D'_{it}$ and the ratio of average annual death counts from 2009 - 2019 compared to corresponding average fatalities over this 11-month period as:

$$Ratio_i = \left( \frac{\sum_{t=2009}^{2019} D_{it}}{\sum_{2009}^{2019} D'_{it}} \right) /11. \tag{A.2}$$

The adjusted mortality count, $D_{it}^{adj}$, was then calculated as:

$$D_{it}^{adj} = D'_{it} \times Ratio_i. \tag{A.3}$$

Thus, the adjusted number of deaths was the actual 11-month (March through December and February) count for the specified year multiplied by the ratio of annual to 11-month (without January) fatality counts over the entire 11-years. This preserves the overall number of deaths from 2009-2019 but removes the volatility in January year-over-year changes.

The baseline number of 2020 year deaths, $D_i^B$, was estimated from regressions of adjusted annual mortality counts on a linear or quadratic time trend:

$$D_{it}^{adj} = \alpha_i TR_t + \epsilon_i, \tag{A.4a}$$

or

$$D_{it}^{adj} = \alpha_i TR_t + \beta_i TRSQ_t + \epsilon_i, \tag{A.4b}$$

where $TR$ is the linear time trend, $TRSQ = TR \times TR$, the last year of the regression model is 2019 and the first year varies between from 2010 and 2015. The use of Poisson models, which had theoretical appeal since the dependent variables were counts, was tested for in preliminary analyses. However, there was no evidence that these improved model fit and so the somewhat easier to interpret results for regression specifications were focused upon.

The decision of whether to use a linear or quadratic trend, as well as the time-period of analysis, was determined by estimating separate models with both types of trends and for all possible starting years between 2009 and 2015, with 2019 always the ending year. The model with the smallest root mean squared error ($RMSE$) for the specified group or cause (or total mortality) was used, with $RMSE_j$ for group/cause $i$ and the $j^{th}$ specification defined as:

$$RMSE_{ij} = \left(\frac{SSE_{ij}}{n_{ij}-p_{ij}}\right)^{0.5}, \tag{A.5}$$

where $n$ was the number of years in the analysis period (ranging from 5 to 11) and $p$ was the number of parameters estimated (two for the linear models and three for the quadratic specifications). This formulation of $RMSE$ penalized shorter panels and quadratic (relative to linear) trend specifications, due to the degrees of freedom correction. The latest starting year was 2015 to insure that at least five years of data were used in the forecasting exercise.

Table A3 shows the models used for all deaths, the 10 population groups, and the 12 causes. The starting analysis year was 2009, 2010, 2013, 2014, and 2015 in 1, 1, 9, 9, and 3 cases respectively. Linear models were used in 6 cases and quadratic specifications in the other 17.

Baseline deaths in 2020 were then estimated from the regression equation with $TR$ and $TRSQ$, where applicable, set to their 2020-year values. Ninety-five percent confidence intervals (95% CI's) were computed using the standard error of the regression predicted value.

In addition to the estimates using the mortality counts adjusted for excess January volatility, estimates were obtained using unadjusted counts. Table A6 shows the model specifications that minimize RMSE in these cases. These are the same specifications as when using the adjusted data in 15 cases, but differ in the remaining 8, two groups (blacks and $\geq$ 65 year olds) and six causes (heart, cancer, diabetes, lower respiratory, influenza/pneumonia, and kidney disease).

*Testing Adequacy of the Prediction Methods*

One way to examine the plausibility of the baseline estimates was to check whether the implied change in 2020 death counts, relative to the period one-year earlier, were in the range of observed prior year-over-year actual death counts. These are often referred to as "baseline changes" and "historical changes" for brevity. To implement this test, historical annual changes, from 2009-2019, were calculated overall and for each group and cause. The average historical change was computed as was the range between the lowest and highest year-over-year changes. Next the baseline change, and the associated 95% CI, were determined from results of the regression procedure discussed above.

Close accordance between the baseline and typical patterns of historical annual changes provided evidence that the estimation models were plausible. Conversely, discordant patterns, such as negative estimated baseline changes, despite a usual historical pattern of

mortality growth, raise questions about the adequacy of the estimation models. Since year-over-year growth also exhibits fairly wide variation over time, it is also useful to compare this range of estimates to the 95% confidence intervals of changes in baseline versus prior year deaths. Table A5 and Figure A1 show the results of these comparisons.

**Excess Death Estimates from Prior Studies**

Estimates of excess deaths and the share attributed to COVID-19 were obtained from four prior studies (3–6) that covered periods from at least the start of the U.S. pandemic through the end of calendar year 2020. Two of these investigations (3,4) built upon methods developed in prior analyses covering shorter periods (7–10), so that problems identified here also apply to those earlier investigations.

These studies used data produced by the Centers for Disease Control and Prevention, as part of their *Provisional Death Counts for Coronavirus Disease (COVID-19)* series (11), as does my analysis of this earlier research. Specifically, deaths in 2019 and 2020 were calculated using CDC Provisional COVID-19 datasets, providing information on weekly deaths in 2019 and 2020 (12,13).

Baseline deaths were computed as the difference between actual and excess fatalities, using information provided in the original investigations when available, but with total deaths calculated using the data sources just described above, when not supplied by the authors.

The methods used in prior research have probably often underestimated the number of counterfactual baseline deaths (expected in the absence of COVID-19) and, as a result, overstated both excess deaths and the share of them attributed to causes other than COVID-19. To demonstrate this, implied changes in baseline deaths were calculated for the four studies (3–6). The baseline year-over-year changes were then compared to historical changes from 2009-2019. For all cases, the 12-month lag period was set to start and end on the same day of the week as in the study analysis timespan. For example, when considering excess deaths from March 1, 2020 through January 2, 2021, the 12-month lag was from March 3, 2019 to January 4, 2020, so that both periods contained the same number of days, started on the first Sunday in March and ended on the first Saturday in January. This procedure ensured that the counts were not affected by either day-of-the week or seasonal differences in death rates.

Two of the studies supplied direct information on baseline deaths (3,6), which could be combined with the lagged mortality counts to provide an implied percentage change, relative to 12-months earlier. These two investigations also supplied 95% CI's for baseline mortality, or for excess deaths from which those on baseline deaths could be calculated. The other two

studies (4,5), provided information on excess deaths but did not indicate either actual or baseline fatality counts. In addition, the Institute for Health Metrics Evaluation (5) analysis did not precisely indicate the period studied. The publication claimed to measure excess deaths from March 2020 through May 10, 2021; however, since it was published on May 13, 2021, data on deaths during the final weeks of this period would not have yet been available. Here, the actual analysis period was assumed to be March 1, 2020 through April 24, 2021. The rationale for this was that data on weekly deaths in the provisional COVID-19 deaths datasets first became available with a 10-14 day lag, so that data through the week ending April 24, 2021 would be the latest available for an analysis conducted approximately 10-days before the final results were published. For these two studies, 2020-2021 deaths counts were calculated using previously released versions of the data set mentioned above (13), designed to roughly correspond to the information available to the authors at the time of their analyses. Specifically, for Rossen et al. (4) the data set was updated on April 21, 2021 and for IHME (5) it was updated on May 5, 2021. If the authors had more recent data, they would have obtained higher counts of both total and excess deaths, but it is not obvious how baseline death counts would have changed.

Rossen et al. (4) did not supply confidence intervals but did provide lower-bounds on the number of excess deaths which, when combined with the data on actual lagged deaths, permitted calculation of an upper-bound on the change in baseline deaths compared to the 12-month lag. IHME (5) supplied only a point estimate on the number of excess deaths and so no confidence intervals or bounds could be placed on the implied change in baseline deaths compared to actual deaths one-year earlier.

There was an extra complication when evaluating the IHME (5) estimates. Since their analysis period covered (approximately) March 1, 2020 through April 24, 2021, the 12-month lag would include roughly March 1, 2019 through April 24, 2020. However, since substantial numbers of COVID-19 deaths occurred in March and April of 2020, including deaths from these two months in the calculation of lagged counts would have incorporated some effects of the COVID-19 pandemic. As an alternative, lagged counts were constructed using the following procedure. First, actual deaths were determined for March 2019 through February 2020, reflecting lagged effects for the first 12 months of the analysis period (March 2020 through February 2021). Second, for March 1 through April 24 of 2021, lagged deaths were constructed by taking the number of actual fatalities from the same dates in 2019 and then multiplying these by 1.0165. The multiplication factor reflected the average annual historical percentage increase in deaths and was included since this portion of then lag period was 24 rather than 12 months prior to the timespan of the original analysis.

Additional adjustments were also needed for Sanmarchi et al. (6). Their stated analysis period was February 26, 2020 through December 31, 2020. However, the number of total and COVID-19 deaths claimed for this timespan did not match those from the CDC provisional data (probably because they used different data sources). After some investigation, I determined that the best match using the CDC data was for 2/27/20 – 12/30/20. Over this period, the provisional death data indicated that there were 2,873,335 deaths, compared to the 2,870,292 reported in the original study, a difference of just 0.1%. Reported deaths with COVID-19 as an underlying cause were also similar (346,036 versus 344,730 in the original study).

A second complication was that Sanmarchi et al. (6) reported cases where COVID-19 was the underlying cause of deaths, whereas most other prior analyses, and the original estimates in this paper, allowed COVID-19 to be either an underlying or a contributory cause of death. To make the Sanmarchi et al. estimates comparable, the underlying cause of death estimates were converted those expected using the broader definition of COVID-19 fatalities. This was done by inflating their original numbers by 10.47%, which reflected the difference between the multiple cause and underlying cause number of COVID-19 deaths during their analysis period in the CDC data.

The implied percentage changes in baseline versus actual 12-month lagged deaths were also computed for three studies (8–10) that preceded but were closely related to Woolf et al. (3) and covered earlier and shorter periods. Using similar methods to those described above, the estimated baseline death rates in these investigations were between 2 and 5 percent lower than actual deaths one year earlier, suggesting an even larger understatement of baseline mortality and greater overstatement of excess deaths in those analyses.

Table A1. Specific Causes of Death Analyzed, with ICD-10 Codes [a]

| Cause of Death (Abbreviation) | ICD-10 Codes |
|---|---|
| COVID-19 (Covid) | U07.1 |
| Diseases of the Heart (Heart) | I00-I09, Ill I13, I20-I51 |
| Malignant Neoplasms (Cancer) | C00-C97 |
| Cerebrovascular Disease (Stroke) | I60-I69 |
| Alzheimer's Disease | G30 |
| Diabetes Mellitus (Diabetes) | E10-E14 |
| Chronic Lower Respiratory Disease (Respiratory) | J40-J47 |
| Influenza or Pneumonia (Flu/Pneumonia) | J09-J18 |
| Nephritis, Nephrotic Syndrome, Nephrosis (Kidney) | N00-N07, N17-N19, N25-N27 |
| Motor Vehicle Accidents (Vehicle) | V02-V89.2 with exceptions [b] |
| Drug Overdoses (Drug) | X40-X44, X60-X64, X85, Y10-Y14 |
| Intentional Self-Harm (Suicide) | U03, X60-X84, Y87.0 |
| Assault (Homicide) | U01-U02, X85-Y09, Y87.1 |

[a] ICD-10 codes refer to the underlying cause of death, except for COVID-19, where they indicate either an underlying or contributing cause.

[b] The full listing of ICD-10 codes for vehicle fatalities is (V02–V04, V09.0, V09.2,V12–V14,V19.0–V19.2,V19.4–V19.6, V20–V79,V80.3–V80.5, V81.0–V81.1,V82.0–V82.1,V83–V86, V87.0–V87.8,V88.0–V88.8, V89.0,V89.2).

Table A2. Year-Over-Year Changes in Monthly Deaths and in January Deaths by Cause [a]

| Month/Cause | Mean | Standard Deviation | Coefficient of Variation | Minimum | Maximum |
|---|---|---|---|---|---|
| Year-Over-Year Changes by Calendar Month | | | | | |
| 1 | 4,128 | 21,672 | 5.25 | -29,095 | 35,721 |
| 2 | 3,578 | 9,523 | 2.66 | -10,027 | 15,858 |
| 3 | 3,876 | 6,389 | 1.65 | -4,394 | 14,235 |
| 4 | 3,230 | 2,990 | 0.93 | -668 | 8,823 |
| 5 | 3,643 | 3,148 | 0.86 | -898 | 8,121 |
| 6 | 3,448 | 2,171 | 0.63 | 1,141 | 6,488 |
| 7 | 3,463 | 1,929 | 0.56 | 576 | 7,578 |
| 8 | 3,480 | 1,831 | 0.53 | -63 | 6,391 |
| 9 | 3,300 | 2,965 | 0.90 | -2,072 | 6,494 |
| 10 | 3,037 | 2,791 | 0.92 | -1,856 | 6,445 |
| 11 | 3,768 | 4,095 | 1.09 | -2,351 | 9,467 |
| 12 | 3,815 | 10,563 | 2.77 | -9,655 | 16,650 |
| Year-Over-Year Changes in January by Cause of Death | | | | | |
| All | 4,128 | 21,672 | 5.25 | -29,095 | 35,721 |
| Heart | 550 | 5,092 | 9.26 | -6,349 | 8,317 |
| Cancer | 309 | 864 | 2.80 | -789 | 1,855 |
| Stroke | 214 | 909 | 4.24 | -1,206 | 1,502 |
| Alzheimer's | 315 | 1,667 | 5.29 | -2,449 | 3,333 |
| Diabetes | 175 | 613 | 3.50 | -661 | 1,055 |
| Lower Respiratory | 258 | 2,376 | 9.22 | -2,941 | 4,206 |
| Influenza/Pneumonia | 129 | 3,705 | 28.70 | -6,444 | 5,239 |
| Kidney | 20 | 387 | 19.42 | -606 | 528 |
| Vehicle | 38 | 163 | 4.29 | -140 | 311 |
| Drugs | 342 | 610 | 1.78 | -583 | 1,652 |
| Suicide | 96 | 167 | 1.74 | -209 | 298 |
| Homicide | 39 | 128 | 3.28 | -160 | 227 |

[a] Table shows summary statistics on year-over-year changes in total monthly deaths and in January deaths by cause for the period 2009-2019.

Table A3. Analysis Period and Model Specification Minimizing Root Mean Square Error (RMSE) [a]

| Group/Cause | First Year [b] | Linear/Quadratic [c] |
|---|---|---|
| All | 2013 | Quadratic |
| Male | 2013 | Quadratic |
| Female | 2013 | Quadratic |
| White | 2013 | Quadratic |
| Black | 2014 | Quadratic |
| Hispanic | 2013 | Linear |
| Other | 2014 | Quadratic |
| <25 | 2013 | Quadratic |
| 25-44 | 2014 | Quadratic |
| 45-64 | 2014 | Quadratic |
| ≥65 | 2014 | Linear |
| Heart | 2015 | Linear |
| Cancer | 2015 | Linear |
| Stroke | 2013 | Quadratic |
| Alzheimer's | 2013 | Quadratic |
| Diabetes | 2013 | Quadratic |
| Lower Respiratory | 2014 | Quadratic |
| Influenza/Pneumonia | 2010 | Quadratic |
| Kidney | 2015 | Linear |
| Drugs | 2014 | Quadratic |
| Suicide | 2009 | Linear |
| Vehicle | 2014 | Quadratic |
| Homicide | 2014 | Quadratic |

[a] RMSE = $[SSE/(n\text{-}p)]^{0.5}$, where $SSE$ is the sum of squared errors, $n$ is the number of years (ranging between 4 and 11) and $p$ is the number of parameters estimated (2 in the linear models and 3 in the quadratic specifications).

[b] First year indicates the starting year of the analysis, which ranges from 2009 to 2015. The end year is always 2019.

[c] Models include either a linear trend variable or a quadratic trend, as noted.

Table A4. Observed-Baseline Death Ratios and % of Excess Deaths Attributed to COVID-19

| Group/Cause | Observed-Baseline Death Ratio (95% CI) [a] | % of Excess Deaths Attributed to COVID-19 (95% CI) |
|---|---|---|
| Groups | | |
| All | 1.23 (1.21 to 1.25) | 82.9% (77.0% to 89.7%) |
| Male | 1.24 (1.22 to 1.26) | 81.8% (77.1% to 87.2%) |
| Female | 1.21 (1.19 to 1.24) | 84.2% (76.6% to 93.4%) |
| White [d] | 1.17 (1.15 to 1.19) | 87.3% (78.9% to 97.8%) |
| Black [d] | 1.33 (1.30 to 1.37) | 67.8% (63.4% to 72.9%) |
| Hispanic | 1.51 (1.50 to 1.53) | 86.5% (85.0% to 88.0%) |
| Other [d] | 1.38 (1.36 to 1.39) | 76.4% (74.3% to 78.5%) |
| Age: <25 | 1.15 (1.07 to 1.23) | 12.3% (8.4% to 23.1%) |
| Age: 25-44 | 1.31 (1.24 to 1.40) | 29.1% (24.3% to 36.2%) |
| Age: 45-64 | 1.26 (1.23 to 1.28) | 67.1% (62.6% to 72.4%) |
| Age: ≥ 65 | 1.21 (1.20 to 1.22) | 95.7% (91.3% to 100.5%) |
| Causes | | |
| Heart | 1.05 (1.04 to 1.05) | |
| Cancer | 1.00 (0.99 to 1.00) | |
| Stroke | 1.07 (1.05 to 1.09) | |
| Alzheimer's | 1.16 (1.11 to 1.22) | |
| Diabetes | 1.16 (1.14 to 1.18) | |
| Lower Respiratory | 0.95 (0.92 to 0.99) | |
| Influenza/Pneumonia | 0.96 (0.87 to 1.07) | |
| Kidney | 1.01 (1.00 to 1.02) | |
| Vehicle | 1.17 (1.09 to 1.26) | |
| Drugs | 1.36 (1.19 to 1.59) | |
| Suicide | 0.92 (0.90 to 0.93) | |
| Homicide | 1.45 (1.28 to 1.66) | |

[a] Observed deaths reported using data from Centers for Disease Control and Prevention, as described in METHODS.

[b] Baseline deaths estimated using data from March 2009 through February 2020 and procedures described in METHODS.

[c] Includes deaths with COVID-19 identified as the underlying or a contributing cause of death.

[d] Excludes Hispanics.

Table A5. Average Annual Historical Change in Deaths (2009-2019) and Implied Change in Baseline 2020 vs. Actual 2019 Deaths

| Group/Cause | Average % Change in Deaths Vs. Prior Year, 2009-2019 (minimum & maximum change) [a] | Change in 2020 Baseline Deaths vs. Actual 2019 Deaths (95% CI) [b] |
|---|---|---|
| All | 1.6% (-1.2% to 3.6%) | 0.6% (-1.2% to 2.3%) |
| Male | 2.0% (-0.5% to 3.7%) | 1.2% (-0.3% to 2.7%) |
| Female | 1.3% (-2.0% to 3.8%) | -0.1% (-2.2% to 2.0%) |
| White | 1.2% (-1.8% to 3.3%) | 0.0% (-1.8% to 1.9%) |
| Black | 2.2% (-0.2% to 4.2%) | 1.3% (-1.0% to 3.7%) |
| Hispanic | 4.2% (1.7% to 6.8%) | 4.0% (3.1% to 5.0%) |
| Other | 4.6% (2.0% to 7.6%) | 2.4% (1.3% to 3.5%) |
| <25 | -0.9% (-5.7% to 3.2%) | -6.4% (-12.8% to 0.1%) |
| 25-44 | 2.3% (-2.5% to 8.7%) | -2.3% (-8.3% to 3.8%) |
| 45-64 | 0.9% (-1.8% to 2.0%) | -2.1% (-3.9% to -0.2%) |
| ≥65 | 1.9% (-1.1% to 4.5%) | 2.0% (1.0% to 3.1%) |
| Heart | 1.0% (-1.8% to 3.4%) | 1.5% (0.9% to 2.1%) |
| Cancer | 0.5% (-0.2% to 1.6%) | 0.2% (0.0% to 0.4%) |
| Stroke | 1.7% (-1.8% to 5.6%) | 0.1% (-2.1% to 2.2%) |
| Alzheimer's | 4.6% (-3.3% to 18.5%) | -4.3% (-8.8% to 0.3%) |
| Diabetes | 2.5% (-0.2% to 4.5%) | 3.2% (1.6% to 4.7%) |
| Lower Respiratory | 1.4% (-3.5% to 5.8%) | -2.1% (-5.3% to 1.0%) |
| Influenza/Pneumonia | 0.1% (-20.9% to 14.9%) | -1.4% (-11.2% to 8.4%) |
| Kidney | 0.6% (-9.9% to 3.8%) | 0.9% (-0.2% to 1.9%) |
| Vehicle | 1.0% (-2.6% to 6.8%) | -6.8% (-13.8% to 0.2%) |
| Drugs | 7.0% (-2.8% to 23.0%) | -4.3% (-18.2% to 9.7%) |
| Suicide | 2.6% (-1.1% to 6.1%) | 4.8% (3.0% to 6.6%) |
| Homicide | 1.6% (-4.6% to 11.0%) | -9.1% (-20.8% to 2.6%) |

[a] Actual deaths refer to March 2009 – February 2019 period, using data from Centers for Disease Control and Prevention, as described in METHODS.

[b] Baseline 2020 (March 2020 through February 2021) deaths, calculated as described in METHODS.

Table A6. Analysis Period and Model Specification Minimizing
Root Mean Square Error (RMSE), without Adjustment for Excess January Volatility [a]

| Group/Cause | First Year [b] | Linear/Quadratic [c] |
|---|---|---|
| All | 2013 | Quadratic |
| Male | 2013 | Quadratic |
| Female | 2013 | Quadratic |
| White | 2013 | Quadratic |
| Black [d] | 2011 | Linear |
| Hispanic | 2013 | Linear |
| Other | 2014 | Quadratic |
| <25 | 2013 | Quadratic |
| 25-44 | 2014 | Quadratic |
| 45-64 | 2014 | Quadratic |
| ≥65 [d] | 2009 | Linear |
| Heart [d] | 2013 | Quadratic |
| Cancer [d] | 2014 | Quadratic |
| Stroke | 2013 | Quadratic |
| Alzheimer's | 2013 | Quadratic |
| Diabetes [d] | 2013 | Linear |
| Lower Respiratory [d] | 2013 | Quadratic |
| Influenza/Pneumonia [d] | 2009 | Quadratic |
| Kidney [d] | 2011 | Linear |
| Drugs | 2014 | Quadratic |
| Suicide | 2009 | Linear |
| Vehicle | 2014 | Quadratic |
| Homicide | 2014 | Quadratic |

[a] RMSE = $[SSE/(n\text{-}p)]^{0.5}$, where $SSE$ is the sum of squared errors, $n$ is the number of years (ranging between 4 and 11) and $p$ is the number of parameters estimated (2 in the linear models and 3 in the quadratic specifications).

[b] First year indicates the starting year of the analysis, which ranges from 2009 to 2015. The end year is always 2019.

[c] Models include either a linear trend variable or a quadratic trend, as noted.

[d] Indicates different specification than when excess January volatility is accounted for..

Table A7. Excess and COVID-19 Deaths, March 2020 to February 2021 for Specifications without Adjustments for Excess Volatility in January Mortality

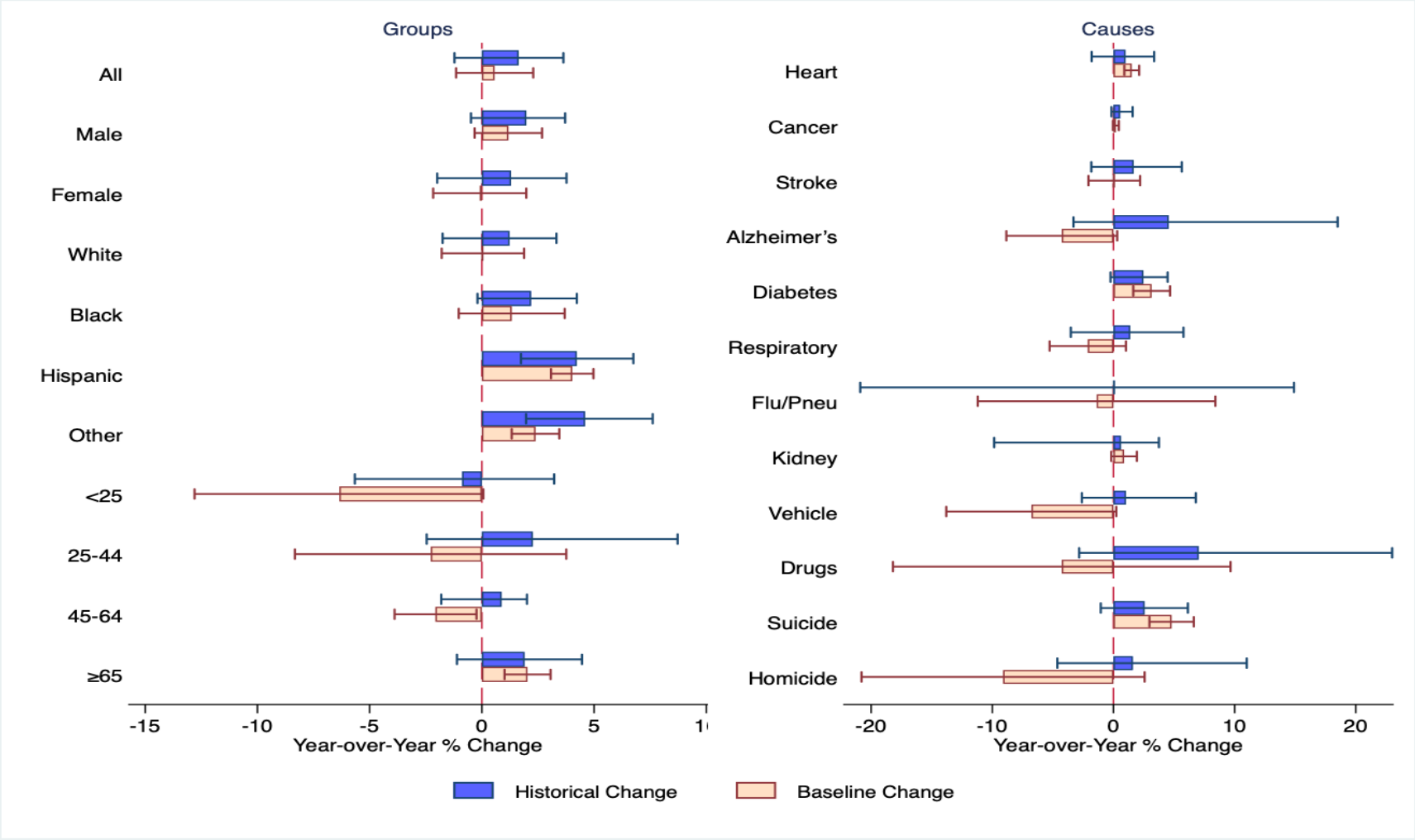
| Group/Cause | Excess Deaths, # (95% CI) [a] | Observed-Expected Death Ratio (95% CI) | % of Excess Deaths Attributed to COVID-19 |
|---|---|---|---|
| All | 665,086 (581,516 to 748,656) | 1.23 (1.20 to 1.27) | 80.9% (71.9% to 92.6%) |
| Male | 366,770 (329,794 to 403,747) | 1.25 (1.22 to 1.28) | 80.3% (73.0% to 89.3%) |
| Female | 298,315 (250,372 to 346,259) | 1.22 (1.18 to 1.26) | 81.6% (70.3% to 97.3%) |
| White [c] | 389,560 (320,659 to 458,462) | 1.18 (1.14 to 1.22) | 84.2% (71.6% to 102.3%) |
| Black [c] | 114,407 (109,785 to 119,029) | 1.32 (1.30 to 1.34) | 70.3% (67.6% to 73.3%) |
| Hispanic | 114,749 (111,994 to 117,505) | 1.52 (1.50 to 1.53) | 86.3% (84.3% to 88.4%) |
| Other [c] | 38,175 (35,902 to 40,448) | 1.38 (1.35 to 1.41) | 75.7% (71.5% to 80.5%) |
| Age: <25 | 8,173 (4,197 to 12,149) | 1.14 (1.07 to 1.23) | 12.6% (8.5% to 24.4%) |
| Age: 25-44 | 43,946 (33,999 to 53,894) | 1.31 (1.23 to 1.41) | 29.2% (23.8% to 37.7%) |
| Age: 45-64 | 135,501 (122,503 to 148,498) | 1.26 (1.23 to 1.29) | 66.8% (61.0% to 73.9%) |
| Age: ≥ 65 | 456,573 (428,504 to 484,642) | 1.21 (1.20 to 1.23) | 95.0% (89.5% to 101.2%) |
| Heart | 41,001 (27,778 to 54,223) | 1.06 (1.04 to 1.08) | |
| Cancer | -1,689 (-4,945 to 1,567) | 1.00 (0.99 to 1.00) | |
| Stroke | 11,336 (7,022 to 15,650) | 1.08 (1.05 to 1.11) | |
| Alzheimer's | 21,441 (14,973 to 27,908) | 1.19 (1.12 to 1.26) | |
| Diabetes | 15,380 (14,330 to 16,430) | 1.17 (1.16 to 1.19) | |
| Lower Respiratory | -7,300 (-14,635 to 35) | 0.95 (0.91 to 1.00) | |
| Influenza/Pneumonia | -1,904 (-10,528 to 6,719) | 0.96 (0.82 to 1.16) | |
| Kidney | -309 (-1,165 to 546) | 0.99 (0.98 to 1.01) | |
| Vehicle | 6,059 (3,385 to 8,734) | 1.17 (1.09 to 1.26) | |
| Drugs | 26,408 (14,812 to 38,004) | 1.36 (1.18 to 1.62) | |
| Suicide | -4,213 (-5,064 to -3,361) | 0.92 (0.90 to 0.93) | |
| Homicide | 7,687 (5,320 to 10,053) | 1.44 (1.27 to 1.66) | |

[a] Excess deaths computed as difference between actual and Baseline deaths.

[b] Includes deaths with COVID-19 identified as the underlying or a contributing cause of death.

[c] Excludes Hispanics.

Figure A1. Historical Average Annual Percent Change in Deaths (2009-2019) and Implied Change in Baseline 2020 vs. Actual 2019 Deaths




Note: Figure shows historical average year-over-year percentage changes in deaths from 2009-2019 and implied percentage change in 2020 baseline deaths versus actual deaths in 2019. Whiskers show the range of year-over-year historical changes in actual changes (left panel) and 95% CI on estimated baseline changes (right panel).

## Appendix References


1. Centers for Disease Control and Prevention. National Center for Health Statistics Mortality Data on CDC WONDER. *CDC Wonder*. 2022;(https://wonder.cdc.gov/mcd.html). (Accessed May 15, 2022)

2. Centers for Disease Control and Prevention. Provisional Mortality Statistics, 2018 through Last Month Request. *Provisional Mortality Statistics, 2018 through Last Month*. (https://wonder.cdc.gov/mcd-icd10-provisional.html). (Accessed May 15, 2022)

3. Woolf SH, Chapman DA, Sabo RT, et al. Excess Deaths From COVID-19 and Other Causes in the US, March 1, 2020, to January 2, 2021. *JAMA*. 2021;

4. Rossen LM, Branum AM, Ahmad FB, et al. Notes from the Field: Update on Excess Deaths Associated with the COVID-19 Pandemic - United States, January 26, 2020-February 27, 2021. *MMWR Morb Mortal Wkly Rep*. 2021;70(15):570–571.

5. Institute for Health Metrics and Evaluation. Estimation of total mortality due to COVID-19. 2021;(http://www.healthdata.org/special-analysis/estimation-excess-mortality-due-covid-19-and-scalars-reported-covid-19-deaths). (Accessed June 2, 2021)

6. Sanmarchi F, Golinelli D, Lenzi J, et al. Exploring the Gap Between Excess Mortality and COVID-19 Deaths in 67 Countries. *JAMA Netw Open*. 2021;4(7):e2117359.

7. Rossen LM, Branum AM, Ahmad FB, et al. Excess Deaths Associated with COVID-19, by Age and Race and Ethnicity - United States, January 26-October 3, 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69(42):1522–1527.

8. Woolf SH, Chapman DA, Sabo RT, et al. Excess Deaths From COVID-19 and Other Causes, March-April 2020. *JAMA*. 2020;324(5):510–513.

9. Weinberger DM, Chen J, Cohen T, et al. Estimation of Excess Deaths Associated With the COVID-19 Pandemic in the United States, March to May 2020. *JAMA Intern Med*. 2020;180(10):1336–1344.

10. Woolf SH, Chapman DA, Sabo RT, et al. Excess Deaths From COVID-19 and Other Causes, March-July 2020. *JAMA*. 2020;324(15):1562–1564.

11. National Center for Health Statistics. COVID-19 Death Data and Resources - National Vital Statistics System. 2021;(https://www.cdc.gov/nchs/covid19/covid-19-mortality-data-files.htm). (Accessed March 17, 2021)

12. Centers for Disease Control and Prevention. Weekly Counts of Deaths by State and Select Causes, 2014-2019 (updated 7/22/21). (https://data.cdc.gov/NCHS/Weekly-Counts-of-Deaths-by-State-and-Select-Causes/3yf8-kanr). (Accessed September 15, 2021)

13. Centers for Disease Control and Prevention. Weekly Provisional Counts of Deaths by State and Select Causes, 2020-2021 (updated 11/3/2021). (https://data.cdc.gov/NCHS/Weekly-Provisional-Counts-of-Deaths-by-State-and-S/muzy-jte6). (Accessed November 5, 2021)